REMARKS

Applicants respectfully request reconsideration of the 35 U.S.C. §102(e) rejection of

claims 1-26 as anticipated by U.S. Patent 6,714,909 to Gibbon et al. (hereinafter "**Gibbon et**

al."), and the 35 U.S.C. §103(a) rejection of claims 7 and 16-17 as unpatentable over **Gibbon et**

al..

Gibbon et al. discloses a system and method for automatically indexing and retrieving

multimedia content. The method may include separating a multimedia data stream into audio,

visual and text components, segmenting the audio, visual and text components based on semantic

differences, identifying at least one target speaker using the audio and visual components,

identifying a topic of the multimedia event using the segmented text and topic category models,

generating a summary of the multimedia event based on the audio, visual and text components,

the identified topic and the identified target speaker, and generating a multimedia description of

the multimedia event based on the identified target speaker, the identified topic, and the

generated summary.

The lowest level of the hierarchy shown in Fig. 1 of **Gibbon et al.** is level 102, which

includes the continuous multimedia data stream, consisting of audio, video and text. As

disclosed in column 3, lines 41-55:

> With the audio, video and text separated as shown 102, linear
> information retrieval is possible. The audio, video and text are
> synchronized in time. Text may be from closed caption provided
> by a media provider or generated by the automatic speech
> recognition engine. If text originates from closed captioning, time

alignment between the audio and text needs to be performed. At the next level, commercials are separated 104. The remaining portion is the newscast 106. The news is then segmented into the anchorperson's speech 108 and the speech from others 110. The intention of this step is to use a detected anchor's identity to hypothesize a set of story boundaries that consequently partition the continuous text into adjacent blocks of text. Higher levels of semantic units can then be extracted by grouping the text blocks into individualized news stories 112 and news introductions or summaries 114.

Applicants submit that the present invention relates only to:

①   audio data feature description schemes, based on hierarchical representations of

audio features, as recited in claims 1-18; and

②   audio video data feature collection description schemes, as recited in clams 19-26.

There is no disclosure in the specification or the claims of the instant application of

continuous <u>text</u> being partitioned into adjacent blocks of <u>text</u>, as disclosed in **Gibbon et al**.

The purpose of **Gibbon et al**. is the hierarchical structuring of a program, especially a

new program, using textual information. Constructing and describing audio program hierarchy

without using any textual information, the purpose of the present invention, differs from **Gibbon**

et al.

In news structuring, mentioned by **Gibbon et al.**, the temporal order of the news

segments varies from its original since it pays more attention to a semantic structure of new

program. On the other hand, in the present invention, the temporal order of the audio segments

does not vary throughout the audio program.

In **Gibbon et al.**, audio data is used only for identifying news and commercials. **Gibbon et al.** places special emphasis on text. **Gibbon et al.** teaches how to segment news summaries and news stories by grouping text blocks. (See column 3, lines 54-56). As to the Examiner's analysis regarding claim 2, **Gibbon et al.** only teaches a method for semantic structuring of a single program, especially a news program. **Gibbon et al.** fails to disclose where a plurality of semantically related programs is represented.

As to the Examiner's analysis regarding claims 3-6, in Fig. 2, **Gibbon et al.** teaches only a semantically high level concept (i.e., structuring) especially for news programs, and accompanies content indicates a segment type of a new program. On the other hand, the present invention merely provides a description scheme for generic audio.

T0, T1, T2 ... in Fig 12 are time stamps of text blocks, not time stamps of audio segments. That is, these time stamps indicate segments derived from text analysis. In addition, A (Anchor), D (Detailed news), and C (Commercial) are just labels for each text block. The present invention describes time codes and audio types for corresponding audio segments.

As to the Examiner's analysis regarding claims 8-12, a set of keywords illustrated in Fig. 16-18 correspond to a set of collected news text which is obtained from closed caption or automatic speech recognition engine, and is determined to have higher importance by using keyword histogram. These keywords are only displayed on the news program browser interface

and not described in a description file. The present invention describes keywords in a description

file, with audio type "keyword."

As to the Examiner's analysis regarding claim 13, **Gibbon et al**. provides no information

about multiple channels. Multiple "channels" mentioned in the present invention indicates a

single audio stream which includes multiple synchronized audio data, e.g., a bilingual stream.

As to the Examiner's analysis regarding claim 14-15, **Gibbon et al**. provides no

description for key events and key objects. Textual description in Fig. 16-18 is text describing

news story content which is obtained from closed caption or automatic speech recognition

engine.

As to the Examiner analysis regarding claims 18-26, **Gibbon et al**. mentions description

generation especially for new programs, and does not teach how to provide a single description

related to a certain feature, across plurality of any programs. The present invention teaches audio

video data feature collection description schemes where a single description related to a certain

feature across plurality of any types of programs, which are not covered by **Gibbon et al**.

Thus, the 35 U.S.C. §102(e) rejection and the 103(a) rejections should be withdrawn.

A Notice of Allowance is earnestly solicited.

If, for any reason, it is felt that this application is not now in condition for allowance, the

Examiner is requested to contact Applicants' undersigned attorney at the telephone number

indicated below to arrange for an interview to expedite the disposition of this case.

In the event that this paper is not timely filed, Applicants respectfully petition for an appropriate extension of time. Please charge any fees for such an extension of time and any other fees which may be due with respect to this paper, to Deposit Account No. 01-2340.

Respectfully submitted,

ARMSTRONG, KRATZ, QUINTOS,
HANSON & BROOKS, LLP

William L. Brooks
Attorney for Applicant
Reg. No. 34,129

WLB/nrp
Atty. Docket No. **001615**
Suite 1000
1725 K Street, N.W.
Washington, D.C. 20006
(202) 659-2930

23850
PATENT TRADEMARK OFFICE

Enclosures:

Q:\FLOATERS\WLB\00\001615\Amendment